THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 25, pp. 17428–17439, June 20, 2008

Impact of Phosphorylation and Phosphorylation-null Mutants on the Activity and Deamination Specificity of Activation-induced Cytidine Deaminase*[S]

Received for publication, March 17, 2008 Published, JBC Papers in Press, April 16, 2008, DOI 10.1074/jbc.M802121200

Phuong Pham[‡], Marcus B. Smolka[§], Peter Calabrese[‡], Alice Landolph[‡], Ke Zhang[¶], Huilin Zhou[§], and Myron F. Goodman[‡1]

From the [‡]Departments of Biological Sciences and Chemistry, University of Southern California, Los Angeles, California 90089-2910, the [§]Ludwig Institute for Cancer Research, Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California 92093, and the [¶]Division of Clinical Immunology/Allergy, Department of Medicine, UCLA School of Medicine, Los Angeles, California 90095

Activation-induced cytidine deaminase (AID) initiates somatic hypermutation and class switch recombination in B cells by deaminating C → U on transcribed DNA. Here we analyze the role of phosphorylation and phosphorylation-null mutants on the biochemical behavior of AID, including enzyme specific activity, processivity, deamination spectra, deamination motif specificity, and transcription-dependent deamination in the presence and absence of RPA. We show that a small fraction of recombinant human AID expressed in *Sf*9 insect cells is phosphorylated at previously identified residues Ser[38] and Thr[27] and also at Ser[41] and Ser[43]. S43P AID has been identified in a patient with hyper-IgM immunodeficiency syndrome. Ser-substituted phosphorylation-null mutants (S38A, S41A, S43A, and S43P) exhibit wild type (WT) activity on single-stranded DNA. Deamination of transcribed double-stranded DNA is similar for WT and mutant AID and occurs with or without RPA. Although WT and AID mutants catalyze processive deamination favoring canonical WRC hot spot motifs (where W represents A/T and R is A/G), their deamination spectra differ significantly. The differences between the WT and AID mutants appear to be caused by the replacement of Ser as opposed to an absence of phosphorylation. The spectral differences reflect a marked change in deamination efficiencies in two motifs, GGC and AGC, which are preferred by mutant AID but disfavored by WT AID. Both motifs occur with exceptionally high frequency in human switch regions, suggesting a possible relationship between AID deamination specificity and a loss of antibody diversification.

AID, a B-cell specific protein, is required for tightly regulated mechanisms of Ig antibody diversification, somatic hypermutation, class switch recombination, and gene conversion. SHM[2] is characterized by an exceptionally high mutation rate $\sim 10^{-3}$ to 10^{-4} per base pair per cell division within the V(D)J rearranged Ig genes (1). CSR is a unique region specific event in which the donor switch region upstream of the IgH C_μ exon (coding for an IgM antibody) recombines with one of the switch regions that is 5′ to each of the downstream C exons, such as C_γ, C_α, or C_ϵ, producing IgG, IgA, or IgE antibody isotypes, respectively (2).

Typically, under tight transcriptional control, AID is induced in activated B cells within germinal centers, causing deamination of cytosine residues in the variable and switch regions (V- and S-regions, respectively) of transcribed Ig loci. Deaminated DNA is subsequently replicated or repaired by different cellular repair mechanisms to give rise to diversified isotype-switched and antigen-specific high affinity antibodies (for reviews, see Refs. 3–5).

Biochemical studies have shown that purified native B-cell AID or recombinant AID protein, expressed in *Sf*9 insect cells or *Escherichia coli*, deaminates C residues on ssDNA but not dsDNA, single-stranded RNA, or DNA/RNA hybrid molecules (6–9). AID acts processively on ssDNA and has a distinctive deamination specificity, favoring C targets in WRC hot spot motifs (where W represents A or T, and R is purine) while avoiding SYC cold spots (where S represents C or G and Y is pyrimidine) (10, 11). These intrinsic biochemical properties of AID are likely to be significant determinants of SHM hallmark features in B-cells, favoring mutations at WRC hot spot motifs and broad clonal heterogeneity of mutations within V-regions (12–14).

Transcription is required for both SHM and CSR to occur (15–17). Since AID does not appear to act on dsDNA, transcription probably plays an essential role in generating transient ssDNA, its principal substrate. The ability of AID to deaminate C residues on dsDNA undergoing active transcription has been demonstrated in cells and cell-free assays. For example, ectopically expressed AID can induce mutations on highly transcribed target gene in fibroblasts (18) and *E. coli* (9,

* This work was supported, in whole or in part, by National Institutes of Health Grants ES013192 and R37GM21422 (to M. F. G.) and Grant GM080469. This work was also supported by the Ludwig Institute for Cancer Research (to H. Z. and M. B. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1 and 2.

[1] To whom correspondence should be addressed: Depts. of Biological Sciences and Chemistry, RRI-201, University of Southern California, Los Angeles, CA 90089-2910. Tel.: 213-740-5190; Fax: 213-821-1138; E-mail: mgoodman@usc.edu.

[2] The abbreviations used are: SHM, somatic hypermutation; AID, activation-induced cytidine deaminase; WT, wild type; CSR, class switch recombination; ssDNA, single-stranded DNA; dsDNA, double-stranded DNA; HIGM-2, hyper-IgM syndrome; MI, mutability index; nt, nucleotide(s).

19), and purified AID expressed in various sources has been shown to act on dC targets on linear dsDNA and closed circular dsDNA model substrates when transcribed by a prokaryotic RNA polymerase (7, 10, 11, 20). AID has been shown to interact with human RNA polymerase II (21), although there are currently no biochemical data for AID acting in conjunction with the human transcription machinery to deaminate dsDNA.

Native AID isolated from stimulated primary B-cell nuclei is phosphorylated on Ser^{38} and Tyr^{184} residues (22, 23), where Ser^{38} phosphorylation occurs on about 5–15% of total AID protein (23). Phosphorylation does not appear to be B-cell-specific, since a similar or greater level of Ser^{38}-phosphorylated AID was observed when expressed in 293T and 3T3-NTZ cells (23). Protein kinase A (PKA) was found to interact with AID (22, 24) and appears to be a primary kinase responsible for phosphorylation at Ser^{38} and possibly at Thr^{27} residues *in vivo* and *in vitro* (22, 24).

The potential role of phosphorylation of AID in SHM and CSR was deduced primarily from studies of AID mutations at phosphorylated residues Thr^{27}, Ser^{38}, and Tyr^{184}. A phosphorylation-defective mutant, Y184A, did not affect CSR (22). In contrast, a mutation at S38A reduced CSR dramatically to no more than 20% of the wild type (22, 24) or moderately (35–80%) (23, 25) in *ex vivo* stimulated B-cells. AID S38A caused a significant decrease of SHM in B-cells (23) and SHM and gene conversion in chicken DT40 cells (26). Phosphorylated Ser^{38} was reported to be required for interaction with the 32-kDa subunit of RPA (replication protein A) (27). AID S38A did not interact with RPA; nor did it catalyze deamination of C on dsDNA in a model T7 RNA polymerase transcription system (22, 27).

There are data, however, indicating a lesser role of Ser^{38} phosphorylation in regulating AID activity. Zebrafish AID, which lacks a serine at a position corresponding to human or mouse Ser^{38}, is fully active in supporting SHM and CSR (28, 29) and gene conversion (26). Recombinant AID expressed in *Sf*9 cells and *E. coli* was found to be active on transcribed dsDNA substrates in the absence of RPA (9–11, 30, 31). Thus, although it is clear that a minority of AID molecules are phosphorylated at Ser^{38}, it is considerably less clear how the biochemical behavior of AID is influenced by either the location of phosphorylated residues or whether or not a particular residue is phosphorylated.

The roles of phosphorylation on the properties of AID *in vitro* are amenable to a comparative biochemical analysis using wild type (WT) AID and phosphorylation-defective mutants. Here we have investigated how replacing individual "phosphorylation-active" Ser residues influences AID specific activity, processivity, deamination specificity, transcriptional-dependent deamination, including its interactions with RPA, and mutational spectra. One such mutant, S43P, has been identified in humans diagnosed with hyper-IgM (HIGM-2) syndrome (32), characterized by the absence of CSR. A comparison of deamination motif specificities favored by S43P in relation to WT AID suggests a possible connection of AID deamination specificity with human immunodeficiency disease.

EXPERIMENTAL PROCEDURES

Enzymes and Substrates—Mutant AID proteins (S38A, S38D, S41A, S41D, S43A, and S43P) were constructed by site-directed mutagenesis (QuikChange site-directed mutagenesis kit, Stratagene) using the pAcG2T-AID vector (6, 10) as the template. Recombinant baculoviruses encoding WT and mutant AID were generated according to the recommended protocol (BD Bioscience). WT and mutant GST-AID proteins were expressed and purified as described previously (10, 11), with an additional of Halt phosphatase inhibitor mixture (Pierce) in the lysis buffer. AID proteins were dialyzed in a buffer containing 20 mM Tris-HCl (pH 7.5), 25 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, and 10% glycerol and stored at −80 °C. *E. coli* single-stranded binding protein and recombinant human RPA were overexpressed in *E. coli* and purified according to published protocols (33, 34). T7 RNA polymerase was purchased from Promega, and ultrapure NTP was purchased from Amersham Biosciences. The active, recombinant catalytic subunit of human protein kinase A was purchased from Calbiochem. M13mp2 gapped DNA and M13mp2T7 covalently closed circular dsDNA substrates were prepared as described (10, 11).

Mass Spectrometry Analysis—Approximately 20 μg of purified AID were reduced with 5 mM dithiothreitol in the presence of 1% SDS and subsequently alkylated with 20 mM iodoacetamide. Protein was then precipitated by the addition of 3 volumes of a solution containing 50% acetone, 49.9% ethanol, and 0.1% acetic acid; the sample was kept on ice for 15 min and then centrifuged for 10 min at maximum speed on a bench top centrifuge. The pellet containing precipitated AID protein was resuspended in 40 μl of 8 M urea, 100 mM Tris-HCl, pH 8.0. The suspension was then diluted by adding 160 μl of 150 mM NaCl, 50 mM Tris-HCl, pH 8.0, and the protein was digested with 2 μg of trypsin (Promega) overnight. After digestion, 0.4% trifluoroacetic acid was added, and the peptides were desalted using a 50-mg C18-Sep-Pak cartridge (Waters). Phosphopeptides were purified by IMAC as previously described (35). Purified phosphopeptides were analyzed by microcapillary-liquid chromatography-electrospray ionization-tandem mass spectrometry on a Thermo Finnigan LTQ quadrupole ion trap mass spectrometer, as described (35). For data analysis, the SEQUEST (version 3.4 beta 2) program running on a Sorcerer system (SageN, San Jose, CA) was used for peptide identification. A data base search was performed using a sub-data base consisting of 50 budding yeast proteins in addition to the human AID sequence. The following variable modifications were considered: +80 Da (phosphorylation) for serine, threonine and tyrosine residues; +16 Da (oxidation) for methionine residues. Up to four variable modifications were allowed per peptide, and the peptide mass tolerance used was 3 Da. A semitryptic restriction was applied, and only the top-matched peptides with a probability score above 0.9 were subsequently considered for close inspection. Each MS/MS spectrum that led to a phosphopeptide identification was manually verified to confirm that all significant ions were accounted for and then validated.

Measurements of Deamination-specific Activity on ssDNA—Specific activities of WT AID and phosphorylation mutants were

measured using ^{32}P-labeled 36-nt ssDNA 5′-AGAAA-AGGGGAAAGCAAAGAGGAAAGGTGAGGAGGT-3′.

Reactions were carried out in a buffer containing 50 mM HEPES (pH 7.5), 1 mM dithiothreitol, 10 mM $MgCl_2$ in the presence of 500 fmol of the substrate DNA, 200 fmol of recombinant GST-AID (*Sf*9-expressed) or 6400 fmol of *E. coli* expressed GST-AID, and 20 ng of RNase A. Following incubation at 37 °C for 5 min, the reactions were quenched by a double extraction with phenol/chloroform/isoamyl alcohol (25:24:1), and the deamination product was analyzed as described previously (6). Specific activities were calculated as amount (fmol) of deaminated substrate/min/μg of enzyme.

Deamination Activity of WT and Mutant AID on Transcribed dsDNA—Transcription-dependent AID deamination on the nontranscribed strand of dsDNA was measured using a linear dsDNA, undergoing active transcription by T7 RNA polymerase (10). In a typical reaction (50 μl), GST-AID (7 pmol of *Sf*9-expressed WT or mutant AID or 50 pmol of *E. coli*-expressed AID), RNase A (200 ng), dsDNA substrate (1 pmol), and T7 RNA polymerase (1 μl) in a reaction buffer containing HEPES (50 mM, pH 7.5), dithiothreitol (5 mM), $MgCl_2$ (10 mM), and NTPs (250 μM) were incubated at 37 °C for 30 min. *E. coli* SSB or human RPA (3, 6, or 12 pmol), if present, were added to the reactions as indicated. The reactions were stopped by extracting twice with phenol/chloroform/isoamyl alcohol (25:24:1). The deaminated products were analyzed using a primer elongation-dideoxynucleotide termination assay (10). Thermo Sequenase (U.S. Biochemical Corp.) was used to extend an 18-mer ^{32}P-labeled primer annealed to the target strand in the presence of three dNTPs and either ddATP or ddGTP (80 μM each). The reactions were carried out for seven cycles (95 °C for 30 s, 55 °C for 45 s, 72 °C for 1 min) and terminated by adding an equal volume of stop solution containing 95% formamide and 20 mM EDTA. The reaction products were resolved by 19% polyacrylamide denaturing gel electrophoresis and analyzed by phosphorimaging. Deamination efficiencies were calculated from extension reactions with the ddA mix as a ratio of the band intensity opposite the converted U template site compared with integrated band intensities opposite and beyond the C template site. The efficiencies were also calculated from extension reactions with ddG as a ratio of integrated band intensities beyond the C template site to the integrated band intensities opposite and beyond the C template site.

Mutation Analysis of AID-targeted C Deamination in Vitro—Deamination specificities of WT and mutant AID were measured using the following reaction conditions: 30-μl volume, 50 mM HEPES (pH 7.5), dithiothreitol (1 mM), $MgCl_2$ (10 mM), gapped DNA (500 ng), RNase A (200 ng), and WT or mutant AID (50 to 100 ng of *Sf*9-expressed AID or 1 μg of *E. coli*-expressed AID). Following incubations for 2.5, 5, and 10 min at 37 °C, the reactions were quenched by a double extraction with phenol/chloroform/isoamyl alcohol (25:24:1). Conversions of C → U on the DNA substrate were detected as white or light blue plaques, indicating C → T mutations in a *lacZα* target gene after transfection into uracil glycosylase-deficient (*ung*$^-$) *E. coli*, as previously described (10).

WT and mutant AID deamination spectra were compared using χ^2 tests with the null hypothesis that the distribution of deaminations at the nucleotide positions is the same for two enzymes. We consider a contingency table where the two columns are the WT and mutant AID, the rows are the nucleotide sites, and the numbers in the cells are the deaminations corresponding to that column and row. In order to deal with the problem of small denominators, we only consider those sites for which the sum of the deaminations for the two enzymes is 5 or greater (we have tried other thresholds and reached the same conclusions).

In order to determine whether a motif is deaminated differently by the WT or mutant enzymes, we performed the following test. For each WT or mutant enzyme (S38A, S41A, S43A, or S43P), we calculated the "site mutability index" by dividing the number of deaminations at each site by the total number of deaminations. For the set of sites sharing the motif, we compared the WT index with each of the four mutant AID indexes by the paired Wilcox test. We used all four mutants, since their spectra are similar to each other (but not the WT), and there are a relatively small number of sites for each motif. Since we repeated this test for each of the 16 motifs, the *p* values in Table 2 have been corrected for multiple tests. The reported value is the probability that if 16 independent tests have been performed, then at least one will be as extreme as observed: $c = 1 - (1 - p)^{16}$, where *c* is the correction of *p* from the Wilcox test.

Processivity Deamination Analysis of WT and Mutant AID—An 85-nt ssDNA substrate with two identical hot spot AGC motifs, containing a fluorescein tag located between the two motifs, was used. Deamination reactions (45-μl volume) were carried out in the presence of 15 pmol of ssDNA substrate, 5–15 pmol of WT or mutant GST-AID, and 200 ng of RNase A for 2.5, 5, and 10 min. After treatment with 4 units of uracil DNA glycosylase (New England Biolabs) and alkali, the deamination products were separated on a 16% denaturing polyacrylamide gel, visualized, and quantified using a FX fluorescence scanner (Bio-Rad). Analysis of processive deamination by a single WT and mutant AID enzyme was carried out as described (36, 37). The single deamination rates are approximately equal at the 5′-C and 3′-C target motifs. The correlated double deamination efficiency is calculated as the probability that a single AID molecule deaminates both 5′ and 3′ target motifs on one ssDNA substrate during a single enzyme-DNA encounter.

RESULTS

Human AID Expressed in Sf9 Cells Is Partially Phosphorylated at Ser Residues 38, 41, and 43 and at Thr^{27}—Mass spectral analysis demonstrates that WT recombinant AID expressed in *Sf*9 insect cells is phosphorylated at four predicted protein kinase motifs; two motifs are the previously identified Ser^{38} and Thr^{27} residues (22, 24), and two newly identified residues are at Ser^{41} and Ser^{43} (Fig. 1). Individual peptide fragments of AID showed that just one of the four residues was phosphorylated. Multiple phosphorylations of individual proteins were not detected despite using an isolation technique that strongly favored the detection of multiply phosphorylated peptide fragments (Fig. 1). These three Ser residues 38, 41, and 43 along with Thr^{27} are predicted to be phosphorylation sites for protein kinase PKA or CAMK2, protein kinase Cδ, CK1, and PKA, respectively (Web Scansite 2.0, medium stringency scan (38)).

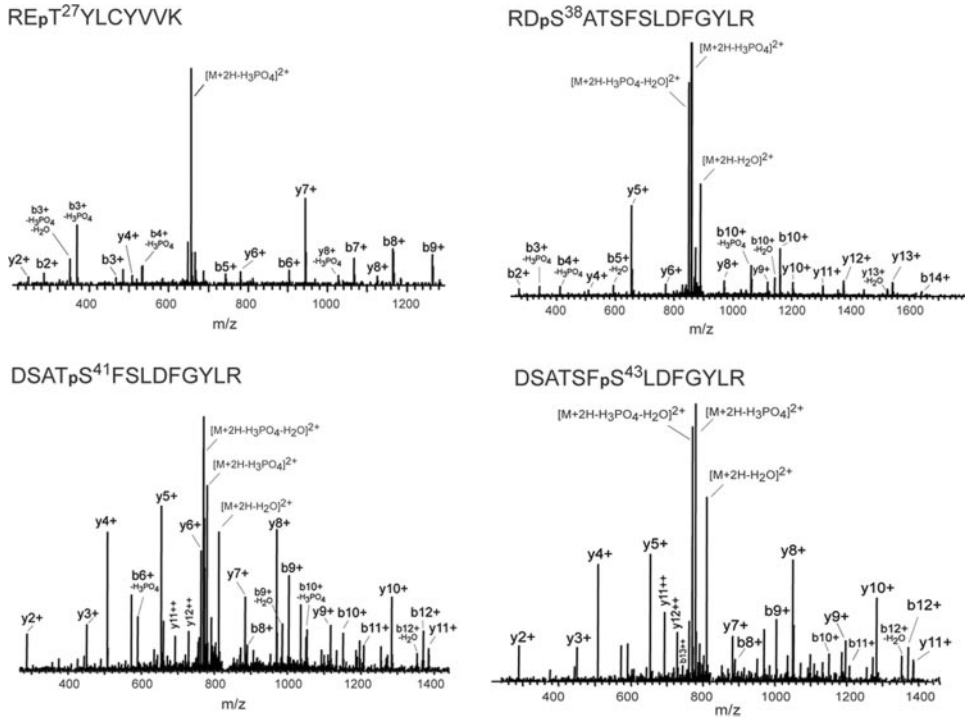

FIGURE 1. **Human AID expressed in *Sf*9 cells is phosphorylated at four sites.** Individual phosphorylation at threonine 27 and serines 38, 41, and 43 was identified by mass spectrometry of tryptic phosphopeptides purified by IMAC. MS/MS spectra of four distinct doubly charged phosphopeptides of AID are shown. In each case, the phosphorylated residue identified is indicated in the sequence (*pS* or *pT*) and is followed by a *number* referring to the amino acid position in the protein. In the spectrum, the peaks corresponding to the detected b-ions (N-terminal fragment) and y-ions (C-terminal fragments) are labeled accordingly. Also, peaks corresponding to the neutral loss of water or phosphate are labeled. Only monophosphorylated peptides were identified.

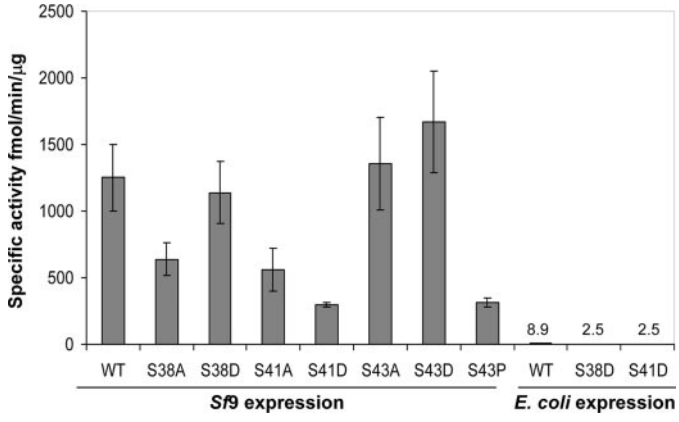

FIGURE 2. **Specific activity of WT and mutant AID proteins expressed in *Sf*9 insect cells or in *E. coli*.** Deamination-specific activities were measured using 36 nt of ^{32}P-labeled ssDNA containing a single C target in an AGC hot spot motif. Reactions were carried out for 5 min at 37 °C in the presence of ssDNA (500 fmol) and recombinant GST-AID (200 fmol) expressed in *Sf*9 cells or with of GST-AID (6400 fmol) expressed in *E. coli*. Due to its extremely low specific activity, the *E. coli*-expressed AID was present at ~32-fold higher concentration than *Sf*9-expressed AID. The *error bars* represent ±1 S.E. from at least three independent measurements.

Based on the amount of protein used in the analysis, we estimate that the phosphorylated species represents less than a few percent of the total AID. The detection of phosphorylated AID required that the enzyme isolation and purification be carried out in the presence of a mixture of phosphatase inhibitors (see "Experimental Procedures"). When phosphatase inhibitors are omitted from the lysis and purification buffers, we were unable to detect any AID phosphorylation by mass spectral analysis (data not shown). Phosphorylation was also not detected for human AID expressed in *E. coli* (data not shown).

Deamination of ssDNA by Phosphorylation-defective AID Mutants—To address whether phosphorylation might regulate AID activity, we constructed single mutants at residues 38, 41, or 43 by replacing Ser with Ala (*i.e.* "phosphorylation null" mutants) or by replacing Ser with Asp, which confers a −1 charge. Deamination measurements were made in the linear range of reaction times and protein concentrations, using the same protein purification protocol (see "Experimental Procedures").

Each AID mutant expressed in insect cells is active on ssDNA, behaving similarly to WT AID within a factor of 3 (Fig. 2). S43A has about the same specific activity as WT AID, whereas S38A and S41A are roughly 2-fold less active. The S38D and S43D mutants are similar to WT AID, whereas S41D and S43P are about 2.5-fold less active. In contrast, human AID expressed in *E. coli* is ~100-fold less active than *Sf*9-expressed AID (Fig. 2). The much lower specific activity of *E. coli*-expressed AID is not likely to be caused by the absence of phosphorylation. Notably, *Sf*9-expressed AID purified in the absence of phosphatase inhibitors showed no detectable amount of phosphorylation but behaved indistinguishably from phosphorylated AID with regard to specific activity and in each biochemical measurement performed below (data not shown). We conclude that the ability of AID to deaminate ssDNA does not depend on phosphorylation; nor does the activity appear to require charged amino acid residues at positions 38, 41, and 43.

Substantial Changes in WT AID Deamination-induced Mutation Spectra Are Caused by Amino Acid Substitutions at Serines 38, 41, and 43—AID was incubated *in vitro* with phage M13mp2, a circular DNA substrate containing a *lacZα* reporter target sequence in a 365-nt single-stranded gapped region (Fig. 3*a*). Deaminations are detected as C → T mutations occurring in DNA isolated from individual phage progeny following transfection of AID-treated DNA into *ung*$^-$ *E. coli* (10). The ratio of mutant phage (clear or light blue plaques) to wild type phage (dark blue plaques) is <5–7%, which ensures that each mutant phage reflects the action of a single AID molecule acting on M13 DNA (Table 1). Conditions that limit M13 DNA targeting to a single AID molecule are satisfied at short (2.5-min) and long (10-min) incubations with AID. We have verified that C → T mutations are absent in DNA isolated from wild type

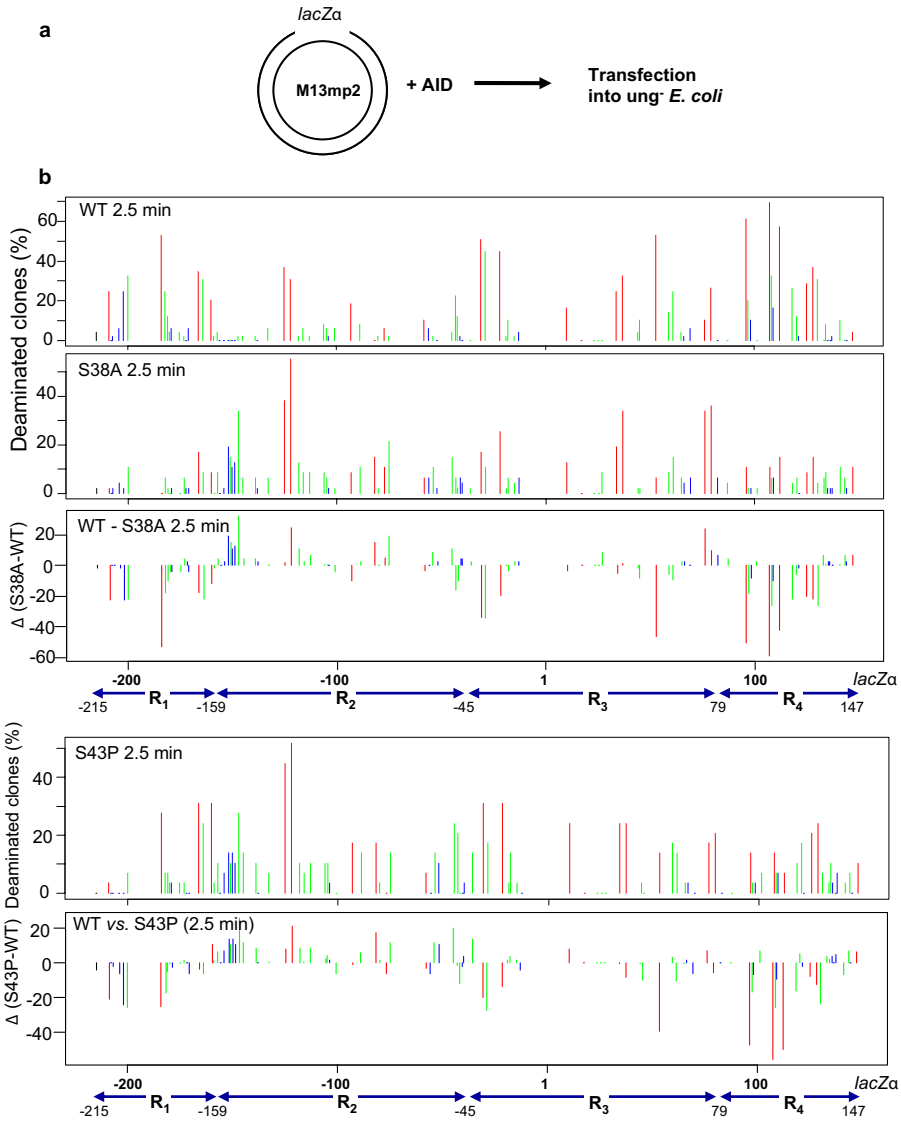

FIGURE 3. **Deamination spectra of WT AID, S38A, and S43P phosphorylation-null mutant AID.** *a*, *sketch* of a protocol to detect C → T mutations resulting from AID-catalyzed C deaminations occurring on individual ssDNA substrate molecules in a gapped M13 DNA construct. The ssDNA gap region is 365 nucleotides long and contains 230 nucleotides of a *lacZα* reporter sequence. Mutations in *lacZα* are identified as *clear* or *light blue plaques*; nonmutated phage appear as *dark blue plaques* (see "Experimental Procedures"). *b*, C → U deamination spectra and the identification of C → T mutations in DNA clones after incubation with WT, S38A, and S43P AID for 2.5 and 10 min were obtained by sequencing about 50 individual mutant clones. Each *colored bar* represents a percentage of mutated phage clones with a C → T mutation at the indicated position on the *lacZα* target sequence (−215 to +149). *Red bars* identify C deaminations occurring in 5′-WRC hot spot motifs, *blue bars* represent 5′-SYC cold spot motifs, and *green bars* represent intermediate motifs (neither WRC nor SYC). WT − S38A and WT − S43P "difference spectra" at 2.5 min and 10 min represent side-by-side comparison of deamination spectra of WT and S38A and WT and S43P, respectively. The *bars* in the difference spectra represent the differences in deamination frequencies at each deaminated site by in the mutant AID spectrum compared with the ones of WT (deamination frequencies in the mutant AID spectrum minus deamination frequencies of WT). *Bars below 0* represent excess deaminations for WT AID, and *bars above 0* denote excess mutant AID deaminations.

phage by direct sequencing of individual DNA clones isolated from dark blue plaques (10).

There are pronounced differences in C → T mutational spectra engendered by differences in the deamination patterns comparing WT AID with S38A and S43P AID mutants, whereas the mutants closely resemble each other (Fig. 3*b*). These differences occur although the WT and mutant enzymes clearly favor deaminations in WRC hot spot motifs and disfavor deaminations in SYC cold spot motifs (Table 2). For both WT and mutant AID, the shorter incubation period (2.5 min) shows a combination of strongly favored deaminations in WRC hot spot motifs (Fig. 3*b* (*red bars*), Table 2, and supplemental Fig. 1) relative to SY*C* cold spot and intermediate deamination motifs (Fig. 3*b*, *blue bars* and *green bars*, respectively). A comparison of WT and mutant AID spatial deamination patterns can be made by dividing the 365-nt ssDNA gapped region into four regions: R_1 (nucleotide positions −215 to −159), R_2 (−158 to −45), R_3 (−44 to +79), and R_4 (+80 to +147). There is a large disparity in the clonal deamination patterns in R_1, R_2, or R_4, comparing mutant S38A and S43P with WT AID (Fig. 3*b*).

To evaluate the spatial differences in deamination quantitatively, we have computed a difference spectrum (Fig. 3*b*, Δ) by subtracting the frequency of deaminated clones in the WT AID spectra from those of S38A and S43P AID at every target C site in the 365-nt gap (Fig. 3*b*). *Lines* extending *below 0* represent excess deaminations for WT AID, and *lines above 0* denote excess S38A and S43P mutant AID deaminations (Fig. 3*b*). The difference spectra reveal a much higher probability of WT AID-catalyzed deaminations, especially in WRC motifs, occurring in R_1, somewhat less so in R_3, and much more so in R_4. In contrast, mutant AID-catalyzed deaminations are clearly favored in R_2 (Fig. 3*b*). A χ^2 test (see "Experimental Procedures") provides strong statistical support that the WT and mutant spectra are significantly different: at 2.5 min, S38A ($p < 2.2 \times 10^{-16}$), S43P ($p < 2.6 \times 10^{-9}$), S41A ($p < 2.2 \times 10^{-16}$), S43A ($p < 2.2 \times 10^{-16}$); at 10 min, S38A ($p < 2.2 \times 10^{-16}$), S43P ($p < 7.4 \times 10^{-4}$), S41A ($p < 4.6 \times 10^{-7}$), S43A ($p < 2.2 \times 10^{-16}$).

For WT AID, there are numerous mutations in WRC motifs (Fig. 3*b*, *red bars*) in the R_1 region, with about 50% of the clones having mutations at C template position −184 and ~18–22% of the clones with mutations at positions −209, −166, and −160 (Fig. 3*b*). Notably, S38A and S43P AID poorly deaminate WRC motifs at sites −209 and −184, which correspond to the two favored hot spot sites in R_1 in the wild type AID spectrum. A second difference is that in the beginning of R_2, a cluster of

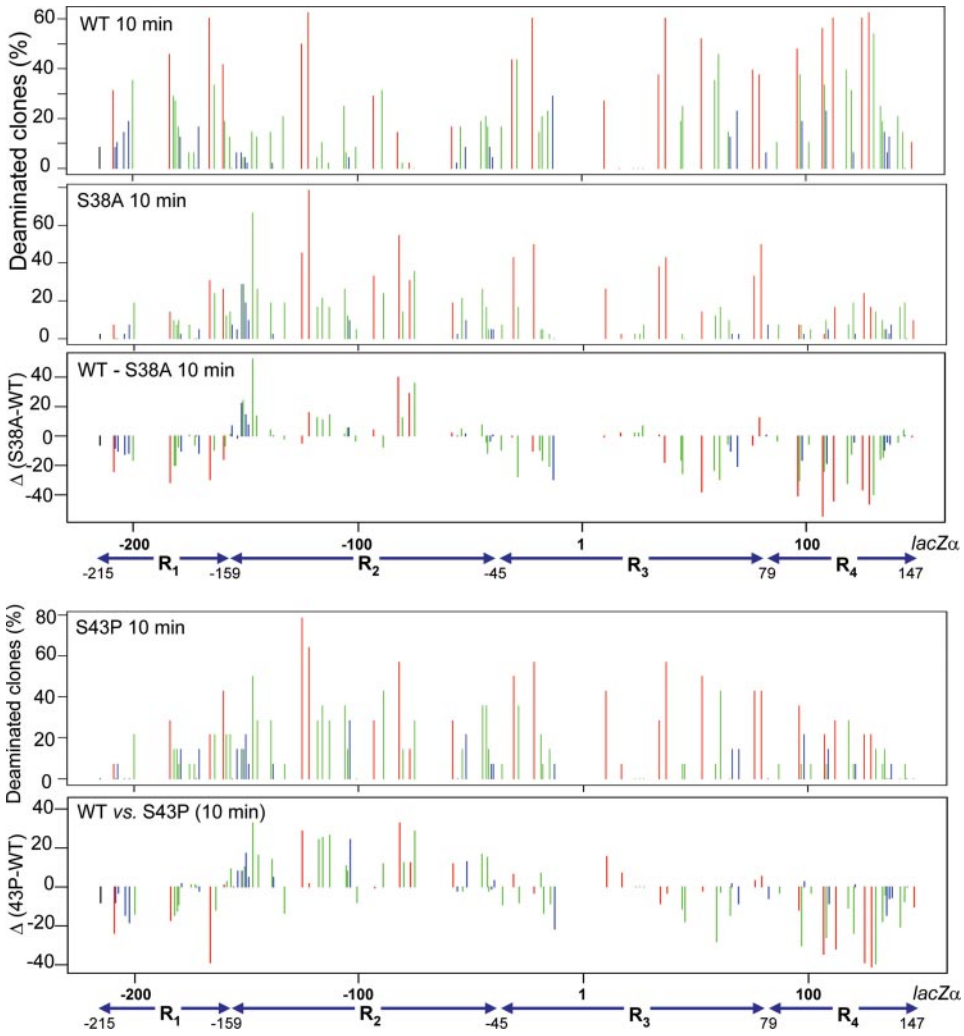

FIGURE 3—*continued*

TABLE 1
Time course analysis of deamination by wild-type and mutant AID on ssDNA gap substrate

	Average number of mutations per clone		Mutant frequency ($\times 10^{-2}$)[a]		
	2.5 min	10 min	2.5 min	5 min	10 min
***Sf*9-expressed**					
WT AID	13	20	3.8	5.1	5.7
S38A	9	16	3.7	4.7	6.1
S41A	11	15	2.7	3.4	4.3
S43A	16	24	4.6	4.7	5.5
S43P	10	20	1.7	2.2	2.5
E. coli-expressed WT AID	ND[b]	17	0.9	1.4	1.5

[a] The background mutant frequency for gapped substrate in the absence of AID is $\sim 0.07 \times 10^{-2}$. *E. coli*-expressed AID was present at 10–20-fold higher levels (1 μg) compared with *Sf*9-expressed AIDs (50–100 ng).
[b] Not determined.

template Cs with SYC cold spots (Fig. 3*b*, *blue bars*) and intermediate motifs (Fig. 3*b*, *green bars*) are deaminated at significant frequencies (10–30%). Deaminations at these sites are virtually absent in the WT AID spectrum (Fig. 3*b*). A major difference between WT and S38A and S43P AID occurs in R_4, where deaminations in WRC motifs (Fig. 3*b*, *red bars*) are present in a much higher fraction in the WT spectra compared with S38A and S43P (Fig. 3*b*). Differences in WT *versus* S38A and S43P AID spectra also occur in R_3 but are far less pronounced.

A mutability index (MI), which denotes the experimentally measured mutation frequency for a particular sequence motif compared with the expected unbiased mutation frequency (39), has been used as a quantitative measure for AID deamination specificity (10). MI values for the 12 trinucleotide motifs for WT, S38A, and S43P AID have been tabulated (Table 2). On average, WT and mutant AID exhibit the same general deamination preferences, strongly favoring hot spots (WRC) over cold spots (SYC), with intermediate motifs in between (Table 2, MI average).

The large disparity in the WT and mutant AID spectra may be caused by differences in the distribution of individual intermediate triplet motifs in the *lacZα* regions and their mutability by WT and mutant AID. For example, in the R_2 region, which strongly favors S38A and S43P over WT AID deaminations, the intermediate motifs CGC and GGC are each represented five and four times, respectively. These two intermediate motifs have higher mutability indexes for S38A and S43P compared with WT AID by factors of ~2.7 ($p = 0.004$) and 3.4 ($p = 0.0001$), respectively (Table 2). In contrast, ACC, which is the intermediate motif most strongly favored by WT over mutant AID by a factor of 3.7 ($p = 0.02$) (Table 2), is absent from R_2.

The spectral differences for WT and mutant AID persist for the longer (10-min) incubation period (Fig. 3*b*). The frequency of deaminated clones at each site increases in an approximately spatially uniform manner in all regions from 2.5 to 10 min for WT and mutant AID. The difference spectra comparing WT with mutant AID for 10-min incubations (Fig. 3*b*) show that WT AID is far more active in R_1, R_3, and R_4, and S38A, S41A, S43A, and S43P AID mutants are more active in R_2, in accord with the 2.5 min incubation data (Fig. 3*b* and supplemental Fig. 1).

A compilation of individual trinucleotide motifs present in Ig switch regions reveals that two of the most highly represented motifs are much more strongly favored for deamination by the phosphorylation-null mutants compared with WT AID. The GGC motif (MI 1.0 S43P, MI 0.23 WT; Table 2) represents 33% of IgA, 19% of IgE, and 15% of IgG4 motifs (Table 3). The AGC motif (MI 2.0 S43P, MI 1.3 WT; Table 2) represents 40% of IgA, 14% of IgE, and 11% of IgG4 (Table 3). Notably, S43P AID is

TABLE 2

Three-nucleotide motif MI for WT and mutant AID (S38A and S43P) at 2.5 min

The MI is defined as the number of times a given trinucleotide motif within a segment of DNA contains a mutation, divided by the number of times the oligonucleotide would be expected to be mutated for a mechanism with no sequence bias.

Motif	MI			
	WT	S38A	S43P	*p* value[a]
Hot spots				
AAC	2.5	1.5	1.6	0.41
AGC	1.3	2.1	2.0	0.06
TAC	3.4	2.0	2.2	0.22
TGC	2.6	2.2	2.6	0.99
MI average (WRC)	2.5	1.9	2.1	
Cold spots				
CCC	0.28	0.52	0.44	0.65
CTC	0.45	0.76	0.53	0.99
GCC	0.05	0.18	0.05	0.07
GTC	0.32	0.67	0.00	0.99
MI average (SYC)	0.28	0.53	0.26	
Intermediates				
ACC	1.2	0.32	0.35	0.02
ATC	1.0	0.63	0.71	0.97
CAC	1.4	0.91	0.96	0.31
CGC	0.5	1.3	1.4	0.004
GAC	0.24	0.5	0.35	0.79
GGC	0.23	0.78	1.0	0.0001
TCC	0.48	1.1	1.1	0.75
TTC	0.53	0.71	0.71	0.46
MI average	0.70	0.78	0.82	

[a] *p* values were calculated by the paired Wilcox test (see "Experimental Procedures").

TABLE 3

Distribution of trinucleotide motifs in Ig switch regions

The number of trinucleotide motifs was calculated for the nontranscribed strand of IgA (mouse germ line IgA switch region, accession number gi:194415), IgE (human IgE switch region, accession number gi:32983), and IgG (human Ig γ-4 switch region, accession number gi:31510).

	Number of motifs		
	IgA	IgE	IgG
Hot spots			
AAC	6	37	24
AGC	79	90	148
TAC	4	10	7
TGC	7	30	83
Cold spots			
CCC	0	49	122
CTC	1	31	106
GCC	2	54	135
GTC	5	17	44
Intermediates			
ACC	4	42	74
ATC	2	20	21
CAC	2	34	97
CGC	3	11	30
GAC	15	62	79
GGC	65	126	196
TCC	2	30	108
TTC	0	22	57

present in a patient with hyper-IgM syndrome characterized by the absence of CSR (32).

Wild Type and Mutant AID Act Processively on ssDNA—The average numbers of C → T mutations were between 9 and 16 for a 2.5-min incubation of the M13 gapped substrate with either WT AID or each of the four AID mutants and were between 15 and 24 for a 10-min incubation (Table 1). At the 10-min time point, roughly two-thirds of the clones acted on by WT AID had more than 10 mutations, with some clones containing 60–70 mutations (supplemental Fig. 2). A similar fraction of clones containing >10 mutations is observed for S38A AID (supplemental Fig. 2), although compared with WT AID, this mutant produces many more clones, about 50%, with 10–20 mutations and a smaller number of clones having >20 mutations. Since more than 90% of the DNA exposed to AID has no mutations, as determined by sequencing DNA isolated from individual blue plaques (see Ref. 10), both WT and mutant AID act processively.

An ssDNA oligonucleotide, containing a fluorescein tag (*F*) located between the two identical AGC hot spot motifs, has been used to detect multiple deaminations occurring on the *same* ssDNA molecule (36, 37) (*i.e.* a single conversion of C → U in either the 3′-motif or 5′-motif or double C → U conversions occurring in both 3′- and 5′-motifs) (Fig. 4*a*). AID acts bidirectionally, resulting in roughly equal deamination frequencies at the 3′ and 5′ targets on the same DNA (Fig. 4*a*).

The "processivity factor" is calculated as the ratio of the observed fraction of double deaminations occurring at both 5′-C and 3′-C on the same ssDNA substrate to the predicted fraction of independent double deaminations (36). Processivity factors between 3.5 and 9 for WT AID and AID mutants (Fig. 4, *a–c*) confirm that processive double deamination is well in excess of double deamination attributable to random encounters with the same ssDNA substrate by two AID molecules. The reduction in processivity factors with time reflects an increased probability of multiple enzyme-DNA encounters when a larger percentage of substrates are deaminated (Fig. 4, *b* and c, 5- and 10-min incubations). The average correlated double deamination efficiency for WT and mutant AID is ~30% (Fig. 4*d*) (*i.e.* 70% of deaminations catalyzed by a single AID molecule occur just once, at either the 3′ or 5′ target motif, and 30% occur processively at both 3′ and 5′ target motifs). The data demonstrate that the WT and mutant forms of AID are processive to about the same extent. Significantly, the processive action of AID seems unaffected by amino acid substitutions that eliminate the possibility of phosphorylation at sites 38, 41, and 43.

AID-catalyzed Transcription-dependent C Deamination Occurs in the Absence of Phosphorylation and RPA—We have examined how phosphorylation of AID affects its ability to deaminate C during active transcription of dsDNA (Fig. 5). T7 RNA polymerase is used to transcribe linear dsDNA having a single AGC hot spot motif on the nontranscribed strand. The deamination of C to U on the nontranscribed strand is detected by measuring the termination of primer elongation caused by incorporation of ddAMP opposite U and ddGMP opposite C (see "Experimental Procedures"). Two primer elongation reactions are performed simultaneously, where ddATP is substituted for dATP (Fig. 5*a*, *ddA lanes*) and ddGTP is substituted for dGTP (Fig. 5*a*, *ddG lanes*). The conversion of C → U is detected by the appearance of a lower doublet termination band corresponding to the incorporation of ddAMP, just *below* the C template position, in the *ddA lanes* (Fig. 5*a*, *arrow, ddA lanes 5*, *7*, *9*, *11*, *13*, *15*) and by the presence of bands migrating past the C template site in the *ddG lanes* (Fig. 5*a*, *bracket* to the *right* of the gels, *ddG lanes 6*, *8*, *10*, *12*, *14*, and *16*). The upper doublet band in the *ddA lanes* occurs by incorporating dGMP opposite residual C moieties not acted on by AID (Fig. 5*a*). AID-catalyzed deamination occurs only when the dsDNA undergoes

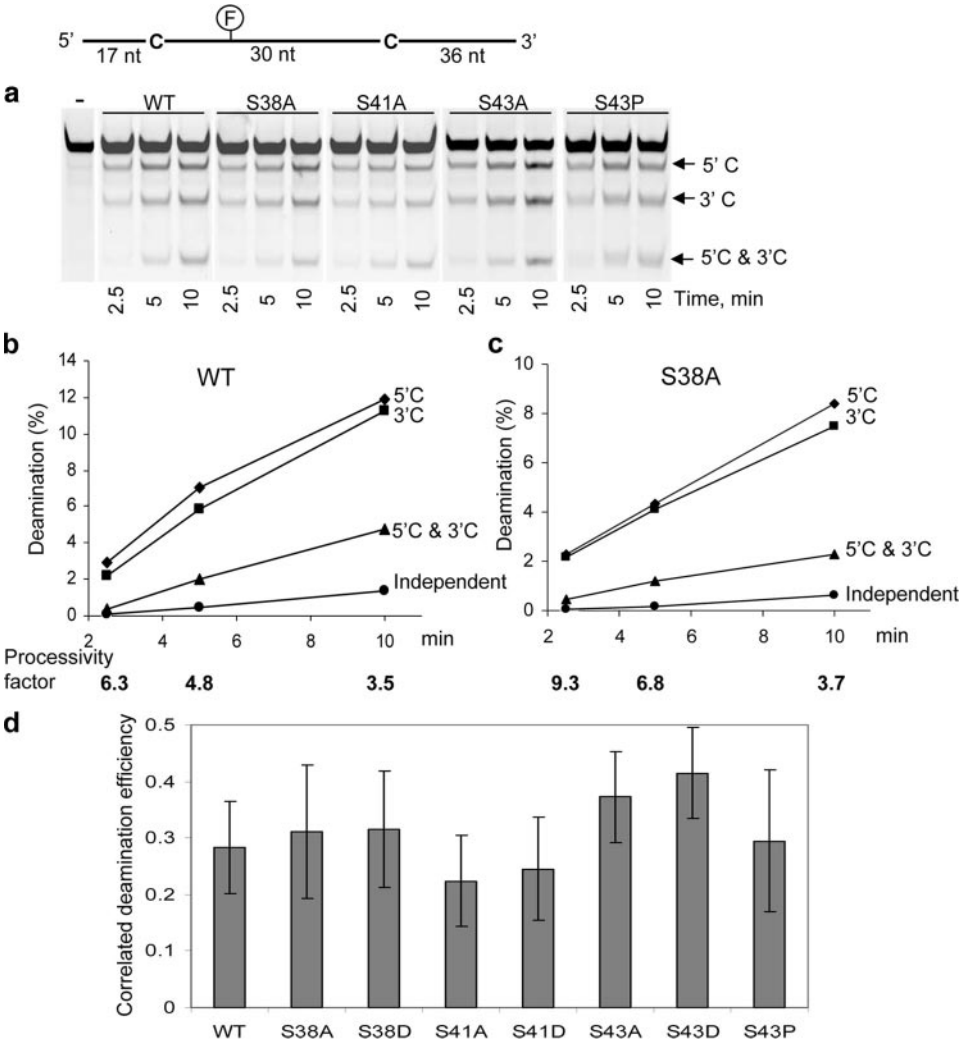

FIGURE 4. **Analysis of processive deamination by WT and mutant AID on the same oligonucleotide ssDNA molecule.** *a*, deamination by WT AID and S38A, S41A, S43A, and S43P phosphorylation-null AID mutants on an ssDNA substrate with two C targets. An ssDNA substrate (85 nt) with two identical C deamination hot spot motifs (AG*C*), containing a fluorescein (*F*) reporter molecule located between the two motifs (*sketch* at the *top*), is used to measure single and double deaminations occurring on an individual ssDNA substrate. Single deaminations of the 5′-C and 3′-C are detected as fluorescein-labeled 67- and 48-nt fragments, *5′C* and *3′C*, respectively; double deaminations of both Cs on the same molecule are detected as a 30-nt fragment (*5′C* & *3′C*). The integrated gel band intensities are used to determine the fraction of correlated double deaminations, generated by the processive action of a single AID molecule (see "Experimental Procedures"). *b* and *c*, deamination (percentage) of substrates containing single deaminations at either 5′-C (♦) or 3′-C (■) and double deaminations at both 5′-C and 3′-C (▲) obtained from the integrated gel band intensities in *a* for WT AID and for S38A mutant AID. ●, the calculated fraction of independent double deaminations (*Independent*), caused by two molecules of AID attacking each site independently. The processivity factor, defined as the ratio of the observed double deaminations occurring at both 5′-C and 3′-C on an individual ssDNA to the calculated fraction of independent double deaminations (36), is shown at the *bottom* of each *graph*. The processivity factor for both WT and mutant AID is significantly greater than 3, indicating that the majority of double deaminations are caused by a single AID molecule deaminating both target C residues. *d*, calculated average correlated deamination efficiency for WT AID and AID mutants. The "correlated deamination" efficiency is the probability for a single enzyme to perform double correlated deaminations at both 3′-C and 5′-C upon a single encounter with an ssDNA substrate. *Error bars*, ±1 S.E.

active transcription in the presence of T7 RNA polymerase (Fig. 5*a*, lanes 5 - 16), whereas no deamination is detected in the absence of either AID or T7 RNA polymerase (Fig. 5*a*, *lanes 1–4*).

The activity of human WT AID expressed in *Sf*9 cells is either similar or slightly less when compared with Ser → Ala phosphorylation-null mutants or with Ser → Asp (−1 negatively charged) mutants (Fig. 5*b*). The activity of the HIGM-2 S43P mutant is indistinguishable from WT AID (Fig. 5*b*). WT AID expressed in *E. coli* requires about a 7-fold higher protein concentration to attain a similar deamination efficiency on transcribed dsDNA as *Sf*9-expressed human AID (Fig. 5, *a* (*lanes 15* and *16*) and *b*). Since the *E. coli*-expressed AID is ~100-fold less active than *Sf*9-expressed AID on ssDNA (Fig. 2), a smaller difference in the relative activities of *Sf*9 and *E. coli* AID on transcribed dsDNA might reflect the possibility that the measured deamination activity in this assay may be dependent on a rate-limiting formation of the transcription bubble needed to provide ssDNA (*i.e.* the nontranscribed strand) as a substrate for AID.

AID-catalyzed transcription-dependent deamination on linear dsDNA occurs in the absence of RPA (Fig. 5). Deamination is inhibited when RPA is present at substoichiometric and higher levels (3–12 pmol) compared with AID (7 pmol) (Fig. 6*a*). We observed ~35–70% inhibition of transcription-dependent deamination, depending on the level of RPA in the reaction, which is shown by a reduction in the intensity of ddA termination bands (Fig. 6*a*, *U arrows*) and by a concomitant reduction in the intensities of bands migrating beyond the target C site in the presence of ddGTP (Fig. 6*a*, *ddG lanes 2*, *4*, *6*, and *8*). RPA inhibits the activities of S38A and S43P AID mutants to about the same extent as WT AID (Fig. 6*b*). A similar inhibitory effect of RPA on WT AID deamination is observed using supercoiled circular DNA, containing a *lacZ* target gene containing numerous hot spot, cold spot, and intermediate target motifs (Table 4).

We investigated the effect of human RPA and *E. coli* SSB on AID activity on a linear DNA oligonucleotide (36-mer, Fig. 6*c*) and gapped (M13 DNA, Fig. 6*d*) substrates. Both single-stranded binding proteins exert a significant inhibitory effect on AID activity. Increasing levels of either RPA or SSB strongly inhibit AID activity on the oligomer substrate (Fig. 6*c*). There is a monotonic reduction in C → T mutant frequencies on the gapped substrate, decreasing to background mutation levels (mutant frequency ~0.07–0.1%) at about 40 pmol of RPA or SSB, which corresponds to a saturating concentration of sin-

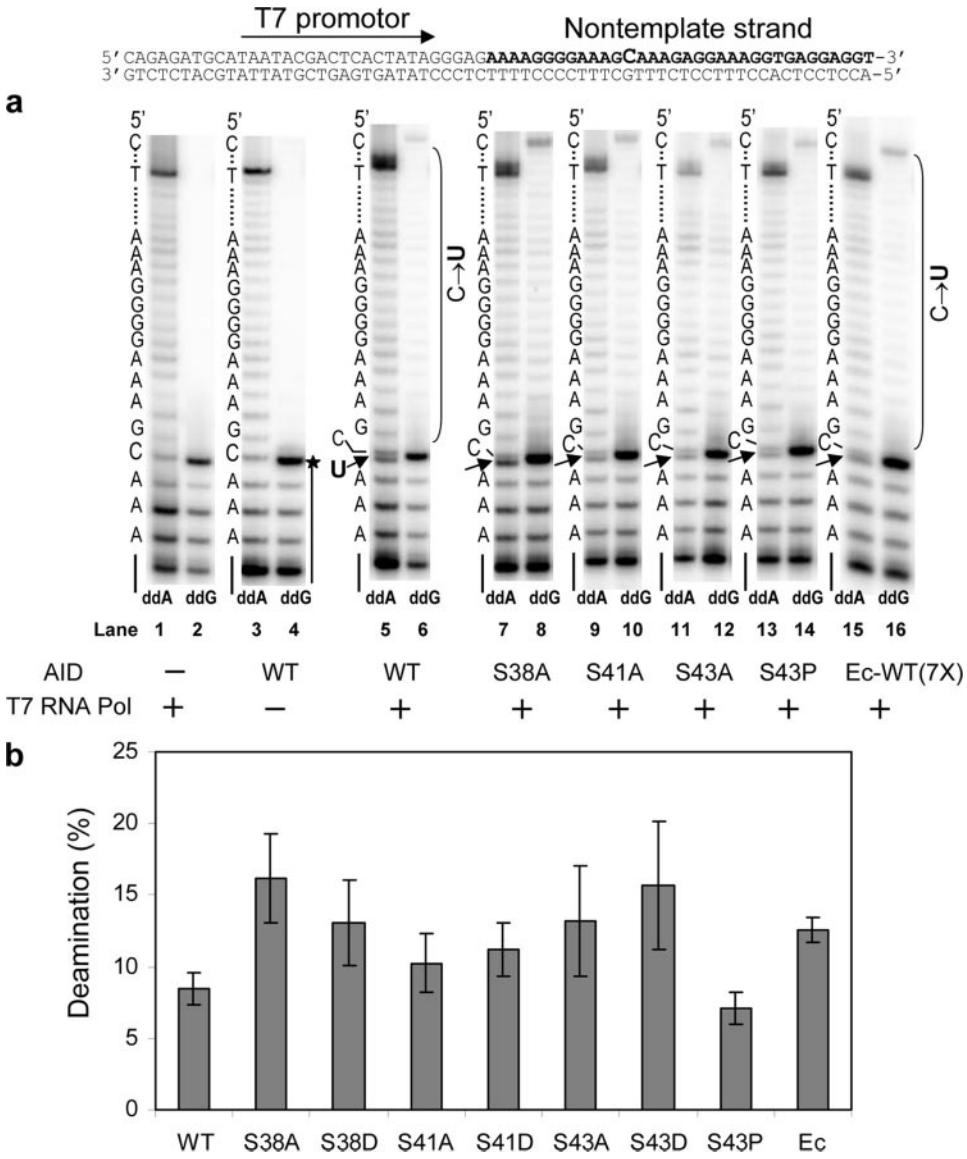

FIGURE 5. **Deamination activity of WT and mutant AID proteins on a transcribed linear dsDNA.** *Top*, dsDNA template with an incorporated T7 promoter for transcription by T7 RNA polymerase. A single target C residue located on the nontranscribed strands is indicated in *boldface type* by a *larger font size*. *a*, polyacrylamide gel electrophoresis data showing C → U deamination on the nontranscribed DNA strand for WT and mutant AID proteins expressed in *Sf*9 cells (*lanes 5–14*) and WT protein, expressed in *E. coli* (*Ec-WT*, *lanes 15* and *16*). A primer elongation dideoxynucleotide termination assay (10) (see "Experimental Procedures") was used to measure the conversion of C → U. Deamination is measured by the appearance of an intense *termination band* just *below* the *template C band* (indicated with an *arrow* on the *left* of each gel) when ddATP is used in place of dATP in the extension reactions (*ddA*, *lanes 5*, *7*, *9*, *11*, *13*, and *15*). Concomitant with the conversion of C to U band in the *ddA lanes* is the presence of bands migrating past the C template site, denoted by the *bracket* to the *right* of the gel, when ddGTP is used instead of dGTP (*ddG lanes 6*, *8*, *10*, *12*, and *14*). Deamination on the dsDNA substrate is not detected in the absence of either AID or T7 RNA polymerase (*lanes 1–4*). The presence of intact C template is observed by the absence of the lower band just below the C template position in the ddA lanes (*lanes 1* and *3*) and by the absence of bands migrating beyond the C template site in the *ddG lanes* (*lanes 2* and *4*). *b*, deamination efficiency of WT and mutant AID on transcribed dsDNA. The deamination reactions were carried out in the presence of 7-fold more (50 pmol) *E. coli*-expressed WT AID (*Ec-WT*) compared with *Sf*9-expressed AID (7 pmol).

gle-stranded binding protein (one RPA trimer or one SSB tetramer to 30-nt ssDNA; see Refs. 40 and 41).

PKA-phosphorylated AID Activity on Single-stranded DNA and on Transcriptionally Activated Double-stranded DNA—Since phosphorylation at Ser[38] has been implicated in efficient CSR *in vivo* (22–24, 27) and in transcription-dependent deamination *in vitro* (22, 27), we have incubated AID with the catalytic subunit of human PKA in the presence of γ-[32]P-labeled dATP, to determine if more extensive phosphorylation at Ser[38], which has a PKA motif, might affect the biochemical properties of AID. We observed a strong phosphorylation signal corresponding to the electrophoretic migration of AID that is dependent on incubation with PKA (Fig. 7*a*, *right-hand gel*), and we estimate that about 50% of the AID is phosphorylated based on the intensities of Coomassie-stained bands corresponding to phosphorylated and nonphosphorylated AID (Fig. 7*a*, *left-hand gel*). Excess phosphorylation has no significant effect on the specific activity of AID acting on ssDNA (Fig. 7*b*). The same conclusion holds for transcription-dependent deamination of dsDNA in the presence or absence of RPA (Fig. 7*c*).

DISCUSSION

AID resides principally in the cytoplasm of activated B cells (42, 43) but is transported into the nucleus, where it catalyzes deamination of C → U on Ig V-gene regions undergoing active transcription. Approximately 10–15% of AID expressed in B cells is reported to undergo phosphorylation at Ser[38] (23). We have found using *Sf*9 insect cells to overexpress human GST-AID that a small fraction of the enzyme is phosphorylated at Ser[38] and Thr[27] and at two residues not reported previously, Ser[41] and Ser[43] (Fig. 1). The data suggest that phosphorylation occurs on just one of these four positions in a single AID molecule. There is considerable interest in understanding the biological roles of phosphorylation of AID, especially with respect to the ability to target transcriptionally active V-regions and switch regions selectively (22, 27). The replacement of Ser[38] with Ala delays the onset of CSR in B-cells (23, 44) and has been reported to abrogate transcription-dependent deamination on linear dsDNA using T7 RNA polymerase *in vitro* (22, 27). The replacement of Ser[43] with Pro engenders hyper-IgM syndrome in humans (32).

In this paper, we aimed to distinguish between how phosphorylation-null mutants *versus* phosphorylation *per se* affects the

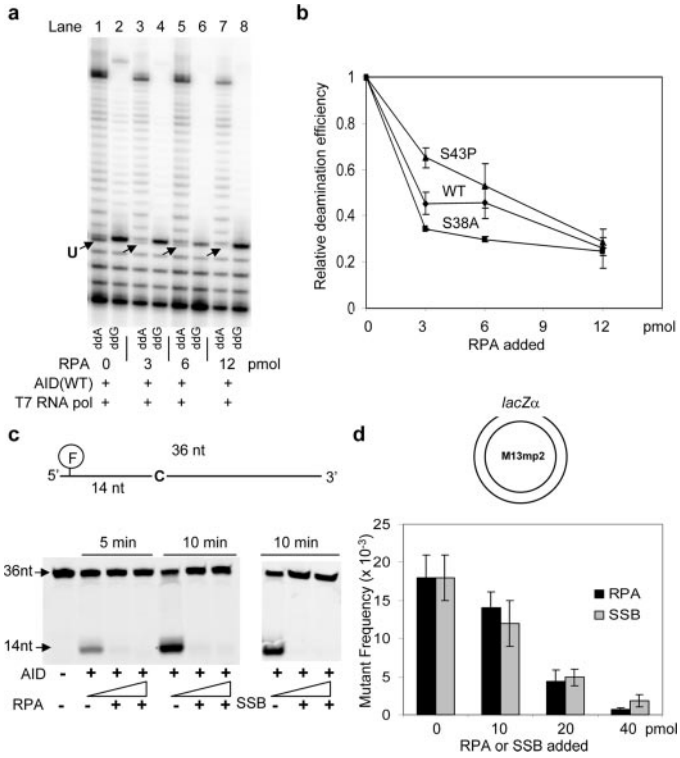

FIGURE 6. **Effect of single-stranded binding proteins (human RPA and *E. coli* SSB) on AID deamination activity.** *a*, deamination of WT AID on the transcribed linear dsDNA in the presence of increasing amounts of RPA. Experiments were carried out as described in the legend to Fig. 5*a*. *b*, relative deamination efficiencies on transcribed dsDNA of WT and mutant S38A and S43P AID in the presence of RPA. Deamination activities in the presence of RPA were normalized to the activity of the enzymes in the absence of RPA. *c*, RPA and SSB completely inhibit activity of WT AID on a fluorescein (*F*)-labeled 36-nt ssDNA substrate with a single hot-spot C (AGC) target. Reactions were carried out in the presence of 7.5 pmol of substrate DNA (255 pmol as nucleotides) and 3 pmol of WT AID for either 5 or 10 min. RPA or SSB, when present, were added at saturating concentration of 10 or 20 pmol. *d*, WT AID-catalyzed deamination on the gapped ssDNA substrate in the presence of RPA and SSB. Deamination activity is monitored by the increase in frequency of mutant (white or light blue color) M13 phage caused by C to U conversion in the *lacZα* target reporter gene. The background mutant frequency for the gapped DNA in the absence of AID is typically 0.7×10^{-3} (10). At 40 pmol, RPA and SSB saturate ssDNA gapped substrates (one RPA trimer or one SSB tetramer to ~30 nt of ssDNA). *Error bars*, S.D. values.

TABLE 4

Effect of RPA on AID activity on transcribed closed circular dsDNA substrate

AID deamination activity was measured using a covalently closed circular dsDNA M13mp2T7 substrate, undergoing active transcription by T7 RNA polymerase (see Ref. 11).

	Mutation frequency ($\times 10^{-3}$)
No AID	0.8 ± 0.3
WT AID	3.7 ± 0.5
WT AID + RPA (60 ng)	4.7 ± 0.9
WT AID + RPA (300 ng)	2.5 ± 0.5
WT AID + RPA (1 μg)	2.2 ± 0.4

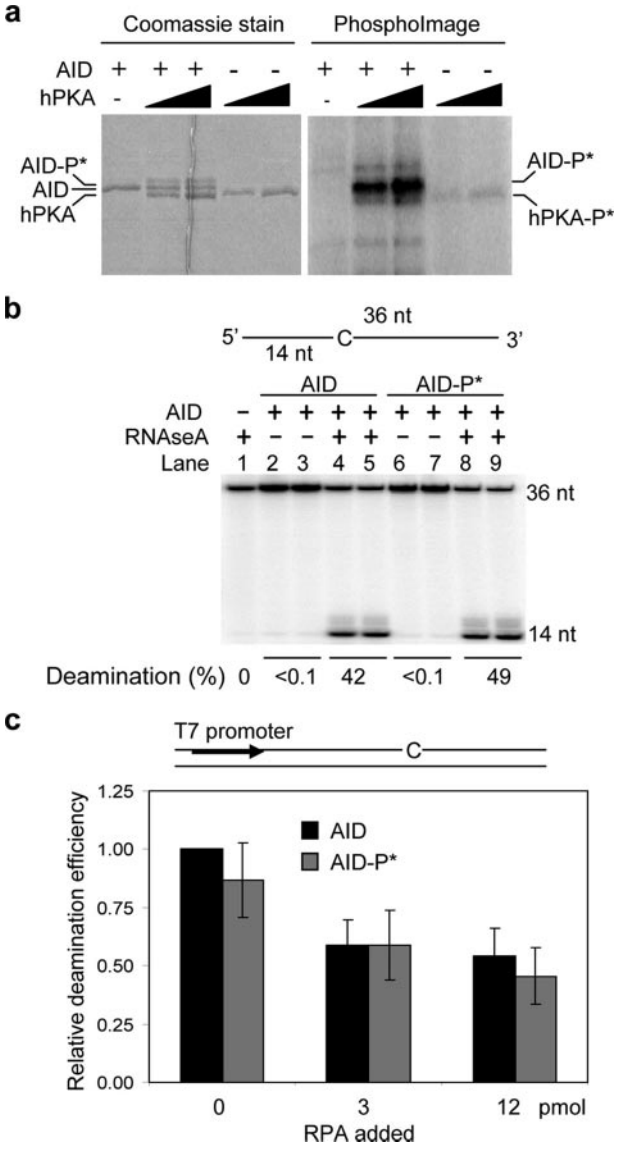

FIGURE 7. **Effect of *in vitro* phosphorylation on AID.** *a*, efficient phosphorylation of human AID by recombinant catalytic subunit of human PKA (*hPKA*). Human PKA and AID were incubated at a 1:2 or 1:1 ratio in the presence of 100 μM ATP and [^{32}P]ATP. About 50% of AID is phosphorylated (*AID-P**) by human PKA (at a 1:1 ratio) as shown by the presence of a slower migrating band on a Coomassie-stained SDS-polyacrylamide gel. The catalytic subunit of hPKA alone is autophosphorylated in the presence of ATP (45). *b*, phosphorylation of AID does not alter the requirement for activation by RNase (*AID-P**; *lanes 6–9*) or change AID activity on ssDNA. *c*, phosphorylation of AID has no stimulating effect on the activity of AID on the transcribed dsDNA even in the presence of RPA. RPA exhibits similar inhibitory effect on both low phosphorylated AID and highly phosphorylated AID (*AID-P**) in a T7-driven transcription deamination assay. The deamination efficiencies were normalized to the activity of low phosphorylated AID in the absence of RPA.

biochemical behavior of AID. The biochemical properties investigated include enzyme specific activity, processivity, deamination spectra, deamination motif specificity, and transcription-dependent deamination, including the influence of RPA. The behavior of insect cell-expressed human AID acting on ssDNA and transcribed dsDNA *in vitro* mimic characteristics of SHM in B-cells, including an approximate simulation of hot spot and cold spot C → T mutational patterns (10). Consequently, a comparative analysis of WT and mutant AID could provide biologically relevant insights into the role of phosphorylation in initiating SHM and CSR.

A comparison of WT AID with phosphorylation-null AID mutants shows major differences in deamination spectra (Fig. 3) and trinucleotide motif specificity (Table 2). However, there appear to be no significant differences in any of the other biochemical properties of AID. The specific activities of WT and mutant *Sf*9-expressed human AID are similar to within a factor of 3 (Fig. 2). Given the inherent variability in enzyme prepara-

tions, these differences are probably not important. Much more important, however, is that phosphorylation of AID at any of the three serine residues is not essential for AID activity. With relevance to human immunodeficiency disease, the S43P mutant associated with HIGM-2 syndrome shows essentially normal activity on ssDNA (Fig. 2) and transcribed dsDNA (Figs. 5 and 6) but generates a significantly different deamination-dependent C → T mutation spectrum compared with WT AID (Fig. 3*b*).

A hallmark property of WT AID is its ability to deaminate ssDNA *in vitro* with high processivity, using a combination of random sliding and jumping modes of action (36, 37). The phosphorylation-null AID mutants are as processive as WT AID; the mutants and WT AID generate about the same average numbers of mutations/clone (Table 1), exhibit similar processivity factors (*e.g.* see Fig. 4*b*, comparing WT AID and S38A mutant AID), and catalyze about the same fraction of correlated double deaminations (Fig. 4*d*).

When T7 RNA polymerase was used to transcribe linear dsDNA (Figs. 5 and 6) or supercoiled circular dsDNA (Table 4), there were no discernible differences in the deamination activity of WT AID compared with phosphorylation-null AID mutants, either in the presence or absence of RPA (Fig. 6). These data are in contrast to a previous report showing that mouse AID obtained from B-cells requires phosphorylation at Ser^{38} and the presence of RPA to deaminate linear dsDNA during transcription by T7 RNA polymerase (22, 27). Although we cannot explain the apparent discrepancy between the two results, our data with phosphorylation-null mutants suggest that neither phosphorylation *per se* nor RPA is needed for transcription-dependent deamination on either linear or circular dsDNA, using GST-tagged human AID expressed either in insect cells (Figs. 5 and 6 and Table 4) or in *E. coli* (Fig. 5) (see also Ref. 9). There is also no apparent increase in transcription-dependent deamination for AID that has been about 50% phosphorylated by the catalytic subunit of PKA, irrespective of whether or not RPA is included in the transcription reaction (Fig. 7).

The presence of a single-stranded binding protein (human RPA or *E. coli* SSB) at a saturating concentration (one RPA trimer or one SSB tetramer per about 30 nt of ssDNA) leads to an almost complete inhibition of AID activity on both short oligonucleotide and gapped ssDNA substrates (Fig. 6, *c* and *d*). We suggest that by binding to ssDNA, RPA and SSB shield templates from deamination by AID. The similar inhibitory effects of RPA and SSB suggest that selective binding of RPA to AID is not occurring, nor is phosphorylation of AID necessary (Fig. 6), at least for the case of human AID expressed in insect cells. The absence of detectable deamination after a relatively lengthy (10 min) incubation time suggests that AID is unable to displace RPA or SSB from ssDNA (Fig. 6*c*). Perhaps RPA might act to protect B cells from a potential "global" mutator activity of AID when DNA becomes transiently single-stranded during normal replication and repair processes.

In contrast to the absence of discernible differences in specific activity, processivity, and transcription-dependent deamination, the replacement of Ser^{38} with Ala or Ser^{43} with Pro has a significant effect on deamination motif specificity and C → T mutation spectra that distinguish the phosphorylation-null mutants from WT AID (Fig. 3*b*). The spectral changes are most pronounced in *lacZα* region R_2 (Fig. 3*b*, Δ difference spectra), where the AID mutants are much more active than WT AID, and in regions R_1 and R_4, where WT AID-catalyzed mutations predominate (Fig. 3*b*, Δ difference spectra). The spectra for all of the phosphorylation-null AID mutants at amino acid residues 38, 41, and 43 are similar to each other (Fig. 3*b* and supplemental Fig. 1) and are substantially distinct from WT AID (p values of $\sim 10^{-7}$).

WT AID and phosphorylation-null AID mutants share the same general motif preferences, where WRC hot spot motifs are generally favored over intermediate motifs and SYC cold spot motifs (Table 2). However, the average MI for WRC hot spot motifs is higher for WT AID (average MI ~2.5) (Table 2) compared with S38A and S43P AID mutants (average MI ~2.0) (Table 2), thus providing a possible explanation for excess WT AID deaminations in regions R_1 and R_4. An analysis of individual DNA clones used to compile the mutational spectra shown in Fig. 3*b* reveals that mutations occur most frequently in clusters (11). We have attributed mutational clustering to a processive bidirectional sliding of AID along ssDNA (10, 11). Mutations initiated in hot spot motifs lead to additional nearby mutations, thus amplifying differences between WT and mutant AID spectra in R_1 and R_4 (Fig. 3*b*, Δ difference spectra).

Favored deaminations occurring in R_2 for the phosphorylation-null AID mutants (Fig. 3*b*) appear to have a more subtle origin that may reflect differences in individual intermediate motif deamination efficiencies. The phosphorylation-null mutants are much more apt than WT AID to attack two intermediate motifs, CGC and GGC (Table 2). The possible relevance of this observation is that CGC and GGC appear five and four times, respectively, in R_2, each having a mutability index favoring mutant over WT deaminations by factors of 2.7 ($p = 0.004$) and 3.4 ($p = 0.0001$), respectively, whereas ACC, which is the favored intermediate motif for WT relative to mutant AID (Table 2, MI ratio ~3.6, $p = 0.02$), is absent from R_2. A clustering of mutations proximal to the intermediate motifs could amplify differences that would favor mutant AID-catalyzed deaminations in R_2.

Studies of AID mutations (T27A and S38A) have led to a suggestion that phosphorylation of AID *in vivo* plays an important role in CSR and SHM. But what that role is has yet to be established. Since phosphorylated AID is preferentially enriched in the chromatin fraction (23), perhaps phosphorylation is needed for transporting AID from the cytoplasm into the nucleus, or to target AID to actively transcribed V-genes and switch regions. Whatever the role turns out to be, it is essential to determine how phosphorylation modulates the biochemical properties of AID. Complicating matters, however, is that zebrafish AID, which has no equivalent Ser^{38} residue, is fully competent in both CSR and SHM (26, 28, 29). Therefore, it is imperative to distinguish between effects attributable to the absence of phosphorylation at a specific Ser amino acid residue *versus* replacement of the Ser residue by a different amino acid (*e.g.* Ala or Pro).

Our data show that the biochemical characteristics of AID, including specific activity, processivity, and transcription-de-

pendent deamination, are not affected either by the absence of phosphorylation on serine residues 38, 41, and 43 or by excess phosphorylation. The replacement of each of these residues, however, results in a significant change in mutant *versus* WT AID deamination spectra. The AID mutants show a high level of deamination notably at two non-hot spot motifs, GGC and CGC, where WT AID is relatively inactive (Table 2).

Switch region DNA sequences contain remarkably high frequencies of GGC motifs. There are 196 GGCs in the S-region of IgG4 and 126 in the S-region of IgE, which are by far the most prevalent triplet motifs in these two switch regions, comprising ~15% in IgG and ~19% in IgE (Table 3). Since WT AID deamination spectra measured on ssDNA and on transcribed dsDNA *in vitro* agree well with C → T SHM spectra in humans (10, 11), by analogy it seems possible, perhaps even likely, that S43A and S43P mutant AID might introduce high levels of U at GGC switch region motifs. Thus, if aberrant recombination were to result from excess deamination targeted to noncanonical switch region motifs, that might contribute to the delay in S38A-initiated CSR in cultured B cells (23) and perhaps to the elimination of CSR in HIGM-2 patients having a S43P mutant form of AID (32).

In a general sense, we have observed that the motifs that WT AID favors *in vitro* are strikingly similar to hot spot motifs of SHM *in vivo* (10), although the enzyme must contend with chromatin, transcription, and assuredly with AID-associated proteins in B cells (3–5). The ability to simulate biochemically several important biological functions of AID suggests that many characteristics of SHM and CSR are determined by its intrinsic properties. Our data suggest the prospect that altered deamination specificity caused by amino acid replacements at Ser^{38}, Ser^{41}, or Ser^{43} could perhaps account for significant deficiencies in SHM and CSR in B-cells.

Acknowledgments—We thank Linda Chelico for preparation of the 85-nt substrate for processivity measurements and Matthew Scharff for a critical reading of the manuscript.

REFERENCES

1. Rajewsky, K., Forster, I., and Cumano, A. (1987) *Science* **238,** 1088–1094
2. Stavnezer, J. (2000) *Curr. Top. Microbiol. Immunol.* **245,** 127–168
3. Goodman, M. F., Scharff, M. D., and Romesberg, F. E. (2007) *Adv. Immunol.* **94,** 127–155
4. Di Noia, J. M., and Neuberger, M. S. (2007) *Annu. Rev. Biochem.* **76,** 1–22
5. Muramatsu, M., Nagaoka, H., Shinkura, R., Begum, N. A., and Honjo, T. (2007) *Adv. Immunol.* **94,** 1–36
6. Bransteitter, R., Pham, P., Scharff, M. D., and Goodman, M. F. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 4102–4107
7. Chaudhuri, J., Tian, M., Khoung, C., Chua, K., Pinaud, E., and Alt, F. W. (2003) *Nature* **421,** 726–730
8. Dickerson, S. K., Market, E., Besmer, E., and Papavasiliou, F. N. (2003) *J. Exp. Med.* **197,** 1291–1296
9. Sohail, A., Klapacz, J., Samaranayake, M., Ullah, A., and Bhagwhat, A. S. (2003) *Nucleic Acids Res.* **31,** 2990–2994
10. Pham, P., Bransteitter, R., Petruska, J., and Goodman, M. F. (2003) *Nature* **423,** 103–107
11. Bransteitter, R., Pham, P., Calabrese, P., and Goodman, M. F. (2004) *J. Biol. Chem.* **279,** 51612–51621
12. Rogozin, I. B., and Kolchanov, N. A. (1992) *Biochim. Biophys. Acta* **1171,** 11–18
13. Smith, D. S., Creadon, G., Jena, P. K., Portanova, J. P., Kotzin, B. L., and Wysocki, L. J. (1996) *J. Immunol.* **156,** 2642–2652
14. Dorner, T., Foster, S. J., Brezinschek, H.-P., and Lipsky, P. E. (1998) *Immunol. Rev.* **162,** 161–171
15. Peters, A., and Storb, U. (1996) *Immunity* **4,** 57–65
16. Fukita, Y., Jacobs, H., and Rajewsky, K. (1998) *Immunity* **9,** 105–114
17. Maizels, N. (1995) *Cell* **83,** 9–12
18. Yoshikawa, K., Okazaki, I. M., Eto, T., Kinoshita, K., Muramatsu, M., Nagaoka, H., and Honjo, T. (2002) *Science* **296,** 2033–2036
19. Ramiro, A. R., Stavropoulos, P., Jankovic, M., and Nussenzweig, M. C. (2003) *Nat. Immunol.* **4,** 452–456
20. Shen, H. M., and Storb, U. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12997–13002
21. Nambu, Y., Sugai, M., Gonda, H., Lee, C., Katakai, T., Agata, Y., Yokota, Y., and Shimizu, A. (2003) *Science* **302,** 2137–2140
22. Basu, U., Chaudhuri, J., Alpert, C., Dutt, S., Ranganath, S., Li, G., Schrum, J. P., Manis, J. P., and Alt, F. W. (2005) *Nature* **438,** 508–511
23. McBride, K. M., Gazumyan, A., Woo, E. M., Barreto, V. M., Robbiani, D. F., Chait, B. T., and Nussenzweig, M. C. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 8798–8803
24. Pasqualucci, L., Kitaura, Y., Gu, H., and Dalla-Favera, R. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 395–400
25. Shinkura, R., Okazaki, I. M., Muto, T., Begum, N. A., and Honjo, T. (2007) *Adv. Exp. Med. Biol.* **596,** 71–81
26. Chatterji, M., Unniraman, S., McBride, K. M., and Schatz, D. G. (2007) *J. Immunol.* **179,** 5274–5280
27. Chaudhuri, J., Khuong, C., and Alt, F. W. (2004) *Nature* **430,** 992–998
28. Barreto, V. M., Pan-Hammarstrom, Q., Zhao, Y., Hammarstrom, L., Misulovin, Z., and Nussenzweig, M. C. (2005) *J. Exp. Med.* **202,** 733–738
29. Wakae, K., Magor, B. G., Saunders, H., Nagaoka, H., Kawamura, A., Kinoshita, K., Honjo, T., and Muramatsu, M. (2006) *Int. Immunol.* **18,** 41–47
30. Shen, H. M., Ratnam, S., and Storb, U. (2005) *Mol. Cell Biol.* **25,** 10815–10821
31. Besmer, E., Market, E., and Papavasiliou, F. N. (2006) *Mol. Cell Biol.* **26,** 4378–4385
32. Zhu, Y., Nonoyama, S., Morio, T., Muramatsu, M., Honjo, T., and Mizutani, S. (2003) *J. Med. Dent. Sci.* **50,** 41–46
33. Lohman, T. M., Green, J. M., and Beyer, R. S. (1986) *Biochemistry* **25,** 21–25
34. Henricksen, L. A., Umbricht, C. B., and Wold, M. S. (1994) *J. Biol. Chem.* **269,** 11121–11132
35. Smolka, M. B., Albuquerque, C. P., Chen, S. H., Schmidt, K. H., Wei, X. X., Kolodner, R. D., and Zhou, H. (2005) *Mol. Cell Proteomics* **4,** 1358–1369
36. Chelico, L., Pham, P., Calabrese, P., and Goodman, M. F. (2006) *Nat. Struct. Mol. Biol.* **13,** 392–399
37. Pham, P., Chelico, L., and Goodman, M. F. (2007) *DNA Repair* (*Amst.*) **6,** 689–692
38. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) *Nucleic Acids Res.* **31,** 3635–3641
39. Shapiro, G. S., Aviszus, K., Murphy, J., and Wysocki, L. J. (2002) *J. Immunol.* **168,** 2302–2306
40. Lohman, T. M., and Ferrari, M. E. (1994) *Annu. Rev. Biochem.* **63,** 527–570
41. Wold, M. S. (1997) *Annu. Rev. Biochem.* **66,** 61–92
42. Rada, C., Jarvis, J. M., and Milstein, C. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7003–7008
43. Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., and Honjo, T. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 1975–1980
44. Basu, U., Chaudhuri, J., Phan, R. T., Datta, A., and Alt, F. W. (2007) *Adv. Exp. Med. Biol.* **596,** 129–137
45. Skalhegg, B. S., and Tasken, K. (2000) *Front Biosci.* **5,** 678–693